

Smoke Screener or Straight Shooter: Detecting Elite Sybil Attacks in User-Review Social Networks

Haizhong Zheng¹ Minhui Xue² Hao Lu¹ Shuang Hao³

Haojin Zhu¹ Xiaohui Liang⁴ Keith Ross²

¹Shanghai Jiao Tong University

²New York University Shanghai

³University of Texas at Dallas

⁴University of Massachusetts Boston

February, 2018



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY



User-Review Social Networks



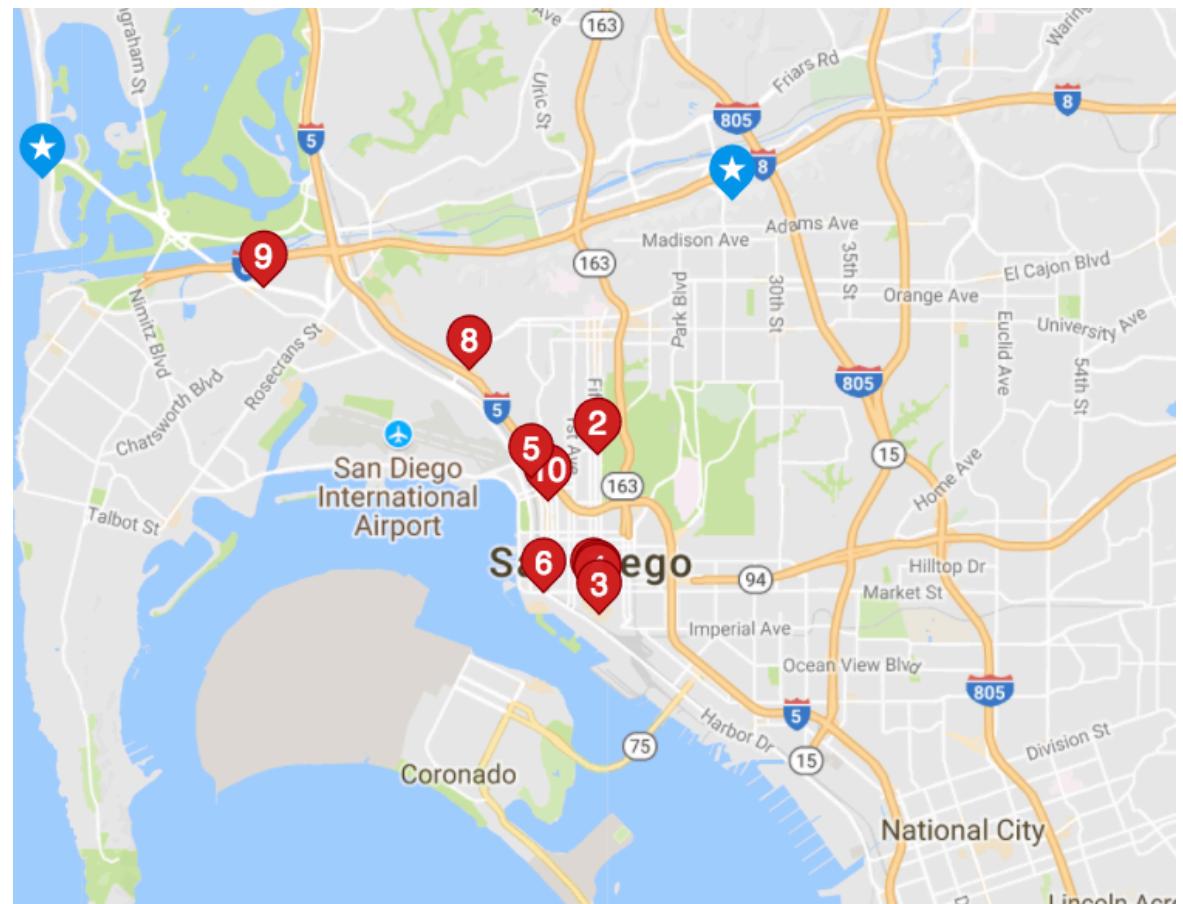
Google My Business



tripadvisor®



Yelp: Search restaurants in San Diego



Reviews in User-Review Social Networks (URSNs)



candybaby918 VIP

口味: 3 环境: 3 服务: 4 人均: 189

[环境]像个精致的大食堂~ 有点开放式餐厅+甜品店的结合~
 [服务]很热情❤️，态度友好，比较温柔，喜欢这种服务，也不会太过
 那个帕尼尼很好吃，份量很大，如果是2姑娘感觉点一份brunch和2杯咖啡就可以，latte很好喝，奶香味浓郁.....

查看全部8张

能不能不吃那么多 VIP

口味: 4 环境: 4 服务: 4

位於downtown的一家brunch or cake店
 dt環境一言難盡 還好這家brunch拯救了我的心情
 店面很大 桌子很多 想來brunch的不需要擔心排隊問題
 喜歡的菜：某牛油果brunch

查看更多

01-26 Bottega Louie

赞(2) 回应 收藏 举报

candybaby918 VIP

口味: 3 环境: 3 服务: 4 人均: 189

[环境]像个精致的大食堂~ 有点开放式餐厅+甜品店的结合~
 [服务]很热情❤️，态度友好，比较温柔，喜欢这种服务，也不会太过
 那个帕尼尼很好吃，份量很大，如果是2姑娘感觉点一份brunch和2杯咖啡就可以，latte很好喝，奶香味浓郁.....

查看全部8张

02-11 Bottega Louie

赞 回应 收藏 举报



BBQ Boss

\$ - Barbeque, Food Stands
 1017 Front St
 San Diego, CA 92101

★★★★★ 2/17/2018

We used BBQ Boss as our caterer for a 40th bday party at Mission Brewery. The food was amazing! What a value. I highly recommend using BBQ Boss to cater your event. They were extremely easy to work with and able to put together a great menu for us. We ordered ribs of the brisket and the mac n cheese. They were both great in quality! Not exaggerating. The meat fell apart and was so well seasoned it didn't need any salt, pepper or bbq sauce. The Mac n cheese had high quality cheeses and a hint of gruyere. I was blown away.

Top notch. Consider them for your next catering event. Excellent value, quality and service. I highly recommend.

Park Place Liquor & Deli

\$ - Delis, Bars
 1000 Park Pl
 Coronado, CA 92118

★★★★★ 1/31/2018

Ordered Online with Yelp

Great sandwiches. Good value. We were headed for the beach and decided to order our picnic for pick up. Sandwiches were very tasty and for \$5-\$8 a piece they were a good deal because they were massive. I highly recommend.

★★★★★ 2/18/2018

1 check-in

I'm definitely a fan! Well worth the wait. The line actually went by pretty quick. I like that you order once you get to the cashier, by the time you get there you already know what to order. After that, you just seat yourself to a vacant table.. and within minutes you will get buzzed to pick up your delicious ribs.
 We got the huge beef ribs with two sides, enough to share for two. Ribs were amazing, along with our colossal Orion rings..
 Will definitely comeback here!

See all photos from Marcel J. for Phil's BBQ



Oceana Champagne Brunch

Reviewed 6 days ago via mobile
 CaseWard, San Diego, California

The food, the service, and the ambience are all amazing! Chef Marcus was very friendly and served up some delicious sunset oysters. Chef Chart served us the BEST sushi! Adri was our server! She always made sure our mimosas were full. Carina, another sever, was... More

A Magical Christmas!

Reviewed 4 weeks ago
 sparklyfromA, Los Angeles, California, United States

After getting married here in May, we returned at Christmas and it was again, an amazing visit! Kyle, one of the managers in reservations, is absolutely THE BEST!!! He always goes out of his way to make sure our visits to the Catamaran is nothing... More

Elevating the SD Experience!

Reviewed January 4, 2018
 Gary M

This was our first trip to San Diego, ideally located on Bay and minutes from the Beach. The views were Spectacular from our suite, I would recommend this resort without hesitation. The dining experience was Awesome, (Tip: sit at the cold bar at Oceana. Chef... More

A peaceful paradise

Reviewed December 4, 2017
 Inimar22, Ashkelon, Israel

Me and my husband are from Israel and we were traveling to USA in October 2017. As a part of our trip, we went to San Diego and took 1 night in this hotel. The service personnel were very kind and professional, always smiling to you. Our room was in the 2nd floor in tower with the garden view. It was great to wake up in the morning with such a wonderful and peaceful view. We loved the nature in that hotel with a big lake surrounds it. It was great to have a dock to boat from this hotel. You can take there a sailboat, kayak or just relax by the ocean as we did. We also liked the pool and jacuzzi which was pretty big. In the evening we went to hotel restaurant "Oceana". The food was great, the service and the whole atmosphere was amazing. With the torches on all over the restaurant, palms, plants and colorful ducks gets out occasionally from the lake by our side, we felt like we were in tropical forest which was so relaxing. We were only sorry that we couldn't stay there more days since we had to continue on our trip. But for those who want to take a few days vacation, this is definitely a wonderful place to stay.
 Show less

Stayed: October 2017, traveled as a couple

Value Cleanliness Service

Ask Inimar22 about Catamaran Resort Hotel and Spa



neal fager 6条评价

★★★★★ 1个月前
 I've only eaten there twice but both times it was Some of the Best Food I've ever eaten in my life!
 I will definitely be going back many times, and probably trying everything on the menu!

The views are also great because its right on ... 详情

顶 分享

Ray Shively 4条评价

★★★★★ 3周前
 The individual that greeted us was a guy who was at best brisk and at worst moderately rude. He did seat us quickly and the seating was quite nice. His manner was horrible.

Food was good not great clam chowder was exceptionally ... 详情

顶 分享

Andrew Sutton 本地向导 23条评价

★★★★★ 1个月前
 Good food and even better view! The food and drink offering, along with the staff, is above average but I typically only eat at World Famous if I am taking a friend or family member that is visiting SD to experience a meal with an ... 详情

顶 分享

Sybil Attacks in URSNs



MarketWatch:

Yelp deems 20% of user reviews 'suspicious'

Published: Sept 27, 2013 8:36 a.m. ET

1,341 320 151 8 95

Aa



Financial Times:

TripAdvisor fined in Italy over false reviews

Online travel website to challenge €500,000 penalty



Phoenix Finance:

Dianping deleted 6 million fake reviews in half a year.



Examples of fake reviews

Which one is the fake review?

Review 1

User level

Star rating of a review

口味4(非常好) 环境4(非常好) 服务4(非常好)

湘菜，对于喜欢吃辣的人来说真是超级赞，不过可能是因为旅游城市的原因，为了适应各地游客的口味，也没有做得特别辣。口味不错，大爱他家的猪肝~名字叫土匪猪肝，然而十分鲜香，入味又没有不新鲜的内脏的那种味道。刁子鱼也不错的，不过我对鱼的品鉴能力有限。湘菜里有名的毛氏红烧肉也做的很好。主食的扬州炒饭饭量很足，总体来说很满意！

喜欢的菜：平锅刁子鱼 土匪猪肝

餐厅特色：干净 免费停车

Pictures of food

05-11 毛家饭店

Review Content

♥ 赞 | 回应 | 收藏 | 不当内容

Review 2

bluefrog蓝蛙(长泰广场店)

上海市浦东新区祖冲之路1239弄长泰广场1座1楼1E27

口味4(非常好) 环境4(非常好) 服务4(非常好)

蓝蛙吃过好多次，喜欢咖啡附送的小饼干，味道和环境也稳定的不错

发表于17-04-15

♥ 赞 | 回应 | 举报

Examples of fake reviews

A fake review posted by the volunteer

The image shows two side-by-side reviews from a platform. A red arrow points from the text 'A fake review posted by the volunteer' to the top review. A blue arrow points from the text 'A real review posted by my friend' to the bottom review.

Top Review (Fake):

- User level: ★★★★ (highlighted with a red box)
- Star rating of a review: ★★★★★ | 口味4(非常好) 环境4(非常好) 服务4(非常好)
- Review content: 湘菜, 对于喜欢吃辣的人来说真是超级赞, 不过可能是因为旅游城市的原因, 为了适应各地游客的口味, 也没有做得特别辣。口味不错, 大爱他家的猪肝~名字叫土匪猪肝, 然而十分鲜香, 入味又没有不新鲜的内脏的那种味道。刁子鱼也不错的, 不过我对鱼的品鉴能力有限。湘菜里有名的毛氏红烧肉也做的很好。主食的扬州炒饭饭量很足, 总体来说很满意!
- Liked dishes: 平锅刁子鱼 土匪猪肝
- Restaurant features: 干净 免费停车
- Pictures of food: Three small images of dishes.
- Rating: 05-11 毛家饭店
- Stamp: FAKE (Large red circular stamp with stars around it)

Bottom Review (Real):

- bluefrog蓝蛙(长泰广场店)
- Address: 海市浦东新区祖冲之路1239弄长泰广场1座1楼1E27
- Rating: ★★★★★
- Review content: 蓝蛙吃过好多次, 喜欢咖啡附送的小饼干, 味道和环境也稳定的不错
- Date: 发表于17-04-15
- Buttons: 赞 | 回应 | 举报

A New Type of Sybils

Elite Sybil users consist of two kinds of users.

1. Sybil accounts creating reviews not belonging to Sybil tasks to **mimic benign users**.
2. **Accounts owned by benign users**, but occasionally fulfilling a Sybil task to get the payments.



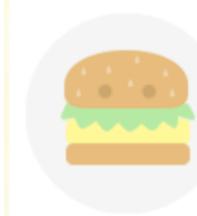
jinranlsj VIP +关注 发私

上海

首页

点评(112)

Elite Sybil User



volencia +关注 发私信

上海

首页

点评(3)

Regular Sybil User

Outlines

01

Elite Sybil Attacks

02

ELISEDET

03

Evaluation & Measurement

04

Conclusion

01

ELITE SYBIL ATTACKS

Sybil Organization Architecture

01

Customers (or Overhyped stores)

Stores want to boost their scores rapidly.

02

Agents

Agents are responsible for accepting tasks from customers and launching campaigns.

03

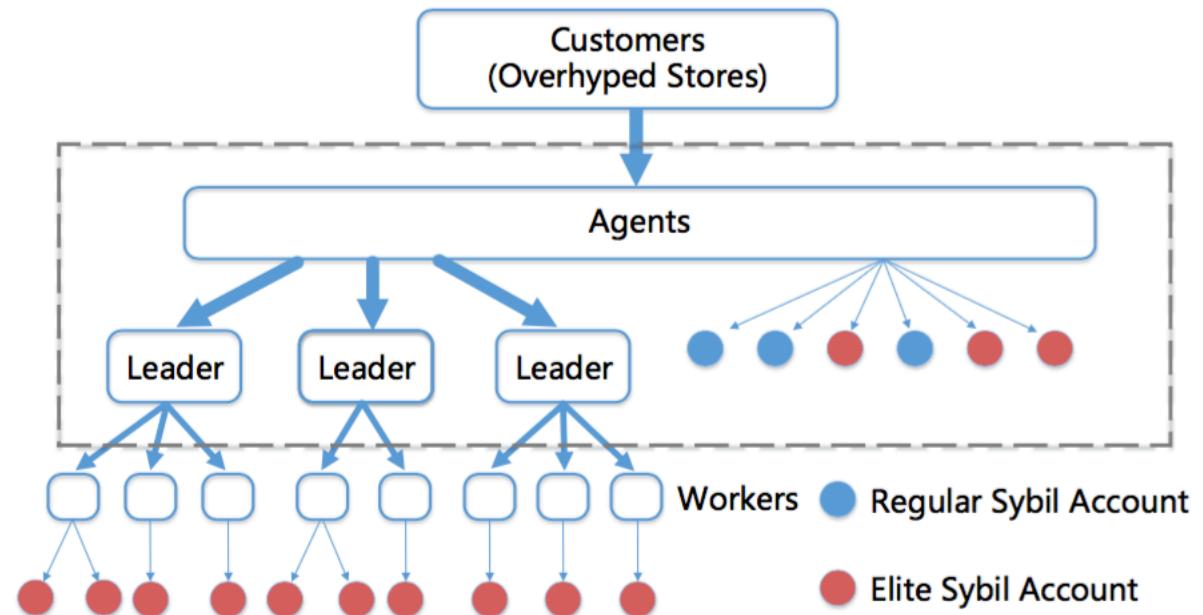
Leaders

Leaders take charge of recruiting workers and distributing tasks to workers.

04

Elite Sybil workers

Elite Sybil workers are internet users post fake reviews for profit.



01

ELITE SYBIL ATTACKS

Sybil Organization Architecture

01

Customers (or Overhyped stores)

Stores want to boost their scores rapidly.

02

Agents

Agents are responsible for accepting tasks from customers and launching campaigns.

03

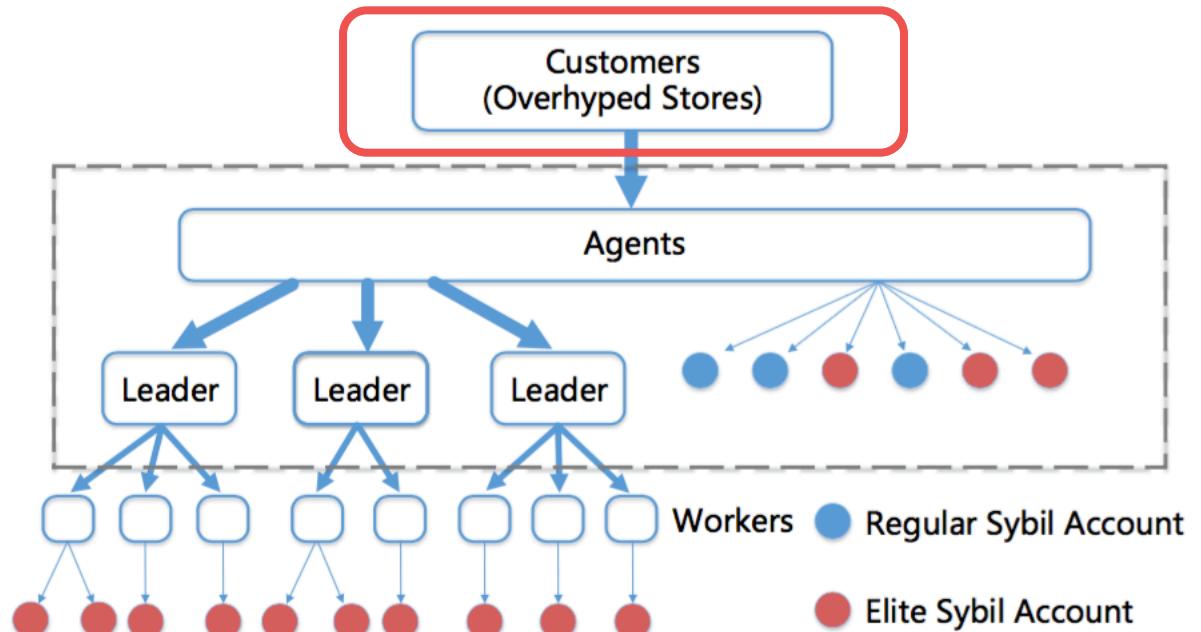
Leaders

Leaders take charge of recruiting workers and distributing tasks to workers.

04

Elite Sybil workers

Elite Sybil workers are internet users post fake reviews for profit.



01

ELITE SYBIL ATTACKS

Sybil Organization Architecture

01

Customers (or Overhyped stores)

Stores want to boost their scores rapidly.

02

Agents

Agents are responsible for accepting tasks from customers and launching campaigns.

03

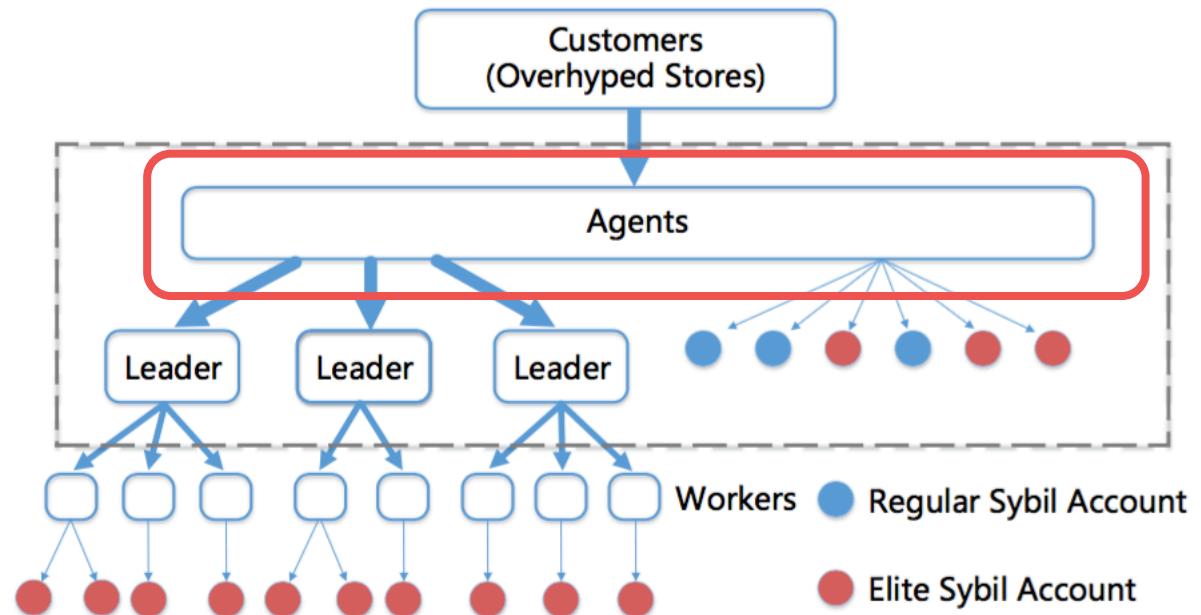
Leaders

Leaders take charge of recruiting workers and distributing tasks to workers.

04

Elite Sybil workers

Elite Sybil workers are internet users post fake reviews for profit.



01

ELITE SYBIL ATTACKS

Sybil Organization Architecture

01

Customers (or Overhyped stores)

Stores want to boost their scores rapidly.

02

Agents

Agents are responsible for accepting tasks from customers and launching campaigns.

03

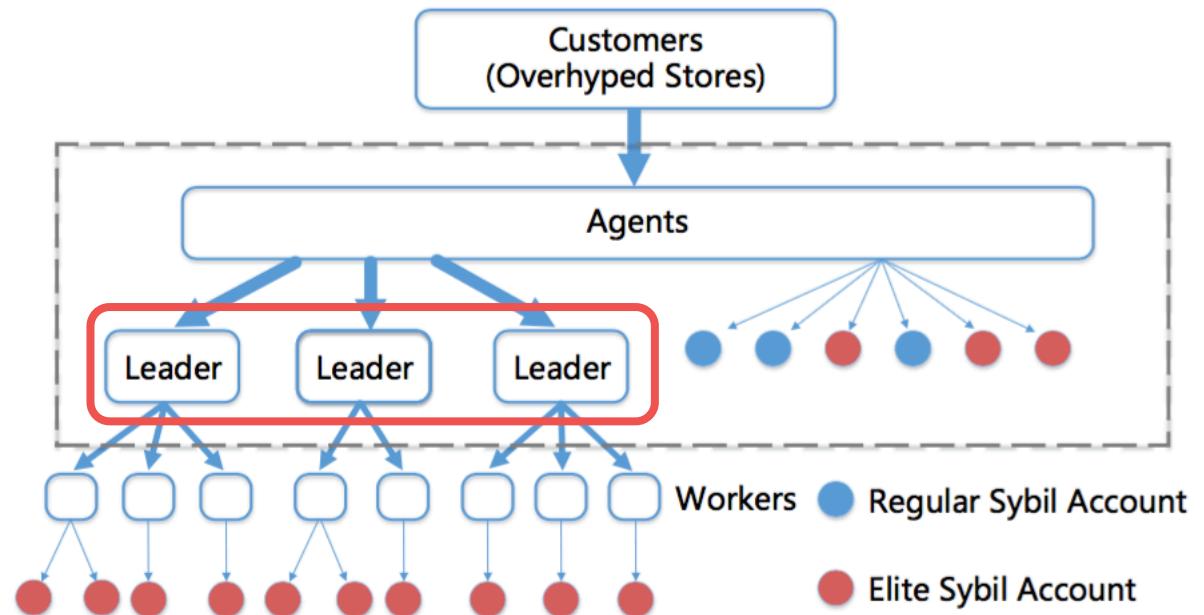
Leaders

Leaders take charge of recruiting workers and distributing tasks to workers.

04

Elite Sybil workers

Elite Sybil workers are internet users post fake reviews for profit.



01

ELITE SYBIL ATTACKS

01

Customers (or Overhyped stores)

Stores want to boost their scores rapidly.

02

Agents

Agents are responsible for accepting tasks from customers and launching campaigns.

03

Leaders

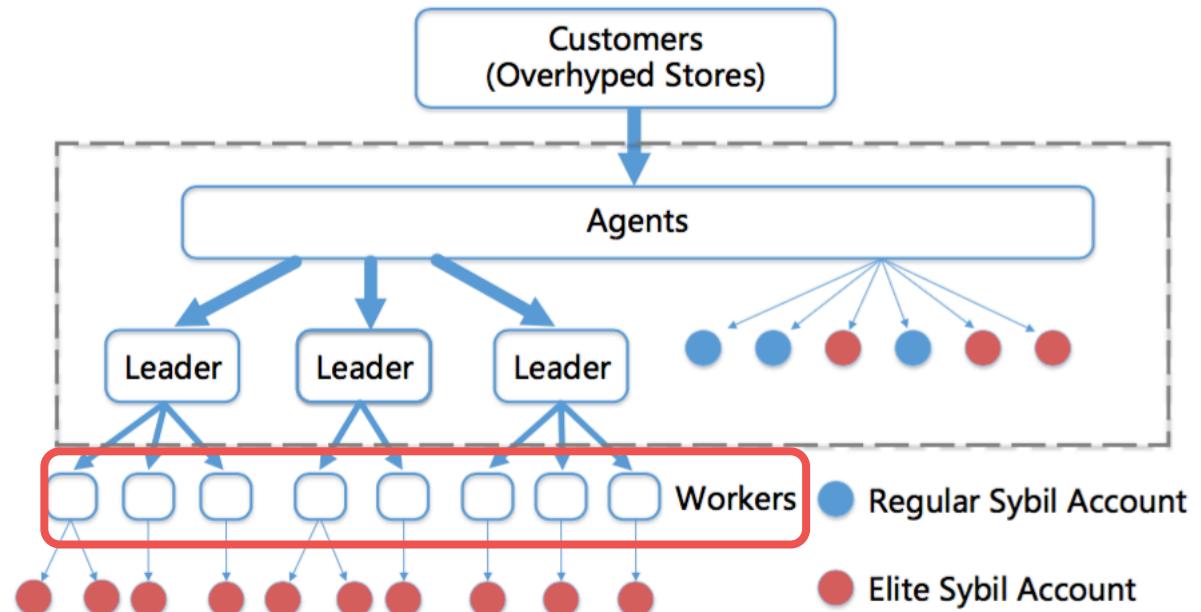
Leaders take charge of recruiting workers and distributing tasks to workers.

04

Elite Sybil workers

Elite Sybil workers are internet users post fake reviews for profit.

Sybil Organization Architecture

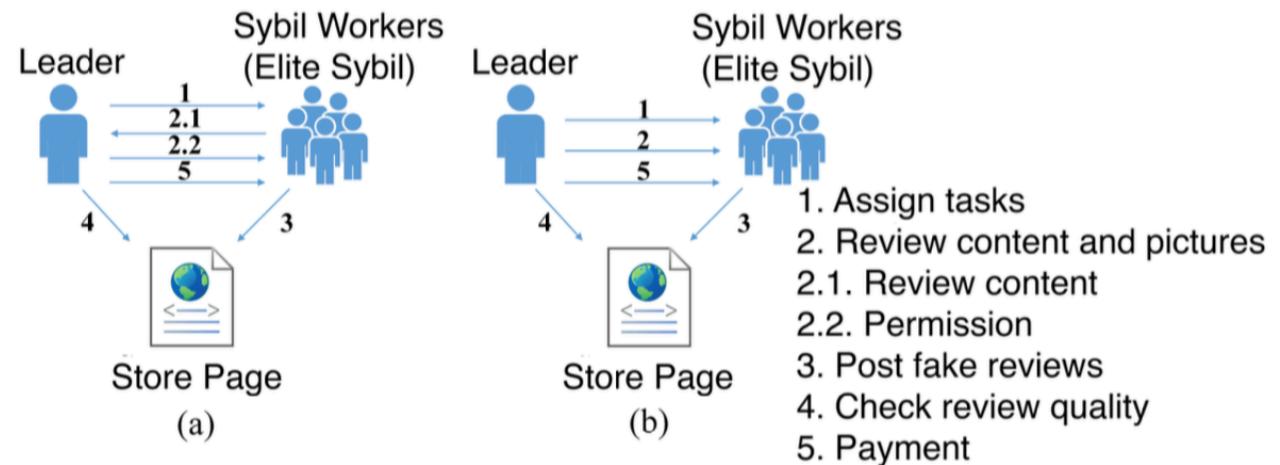


a**Leader-supervised model**

The reviews generated by elite Sybil workers

b**Leader hands-on model**

The content of reviews is provided by leaders.



The quality of fake reviews is strictly controlled!!!

Why do Sybil organizations put so much effort to control the quality of reviews?

The payments vary with ratings of accounts:

TABLE I. HIERARCHICAL REWARDS FOR (ELITE) SYBIL WORKERS

Ratings of Accounts	Rewards per Submission
0-star, 1-star	\$0.30
2-star	\$0.75
3-star	\$1.50
4-star	\$3.74
5-star, 6-star	\$5.98

Weak connectivity of User-Review Social Networks (URSNs)

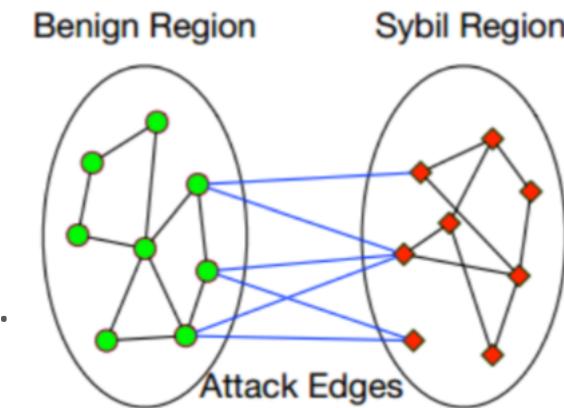
Graph-based detection approaches may not work.

Highly adaptive and professional

Imitate real users in user profiles and review content.
Behavior-based detection approaches may not work.

More active out of Sybil campaigns

Lower percentage of fake reviews in their posts.
Clustering approaches may not work.





User level

★★★★★ | 口味4(非常好) 环境4(非常好) 服务4(非常好) **Star rating of a review**

湘菜，对于喜欢吃辣的人来说真是超级赞，不过可能是因为旅游城市的原因，为了适应各地游客的口味，也没有做得特别辣。口味不错，大爱他家的猪肝~名字叫土匪猪肝，然而十分鲜香，入味又没有不新鲜的内脏的那种味道。刁子鱼也不错的，不过我对鱼的品鉴能力有限。湘菜里有名的毛氏红烧肉也做的很好。主食的扬州炒饭饭量很足，总体来说很满意！

喜欢的菜：平锅刁子鱼 土匪猪肝
餐厅特色：干净 免费停车

Review Content

★★★★  **Pictures of food**

05-11 毛家饭店



Outlines

01

Elite Sybil Attacks

02

ELISEDET

03

Evaluation & Measurement

04

Conclusion



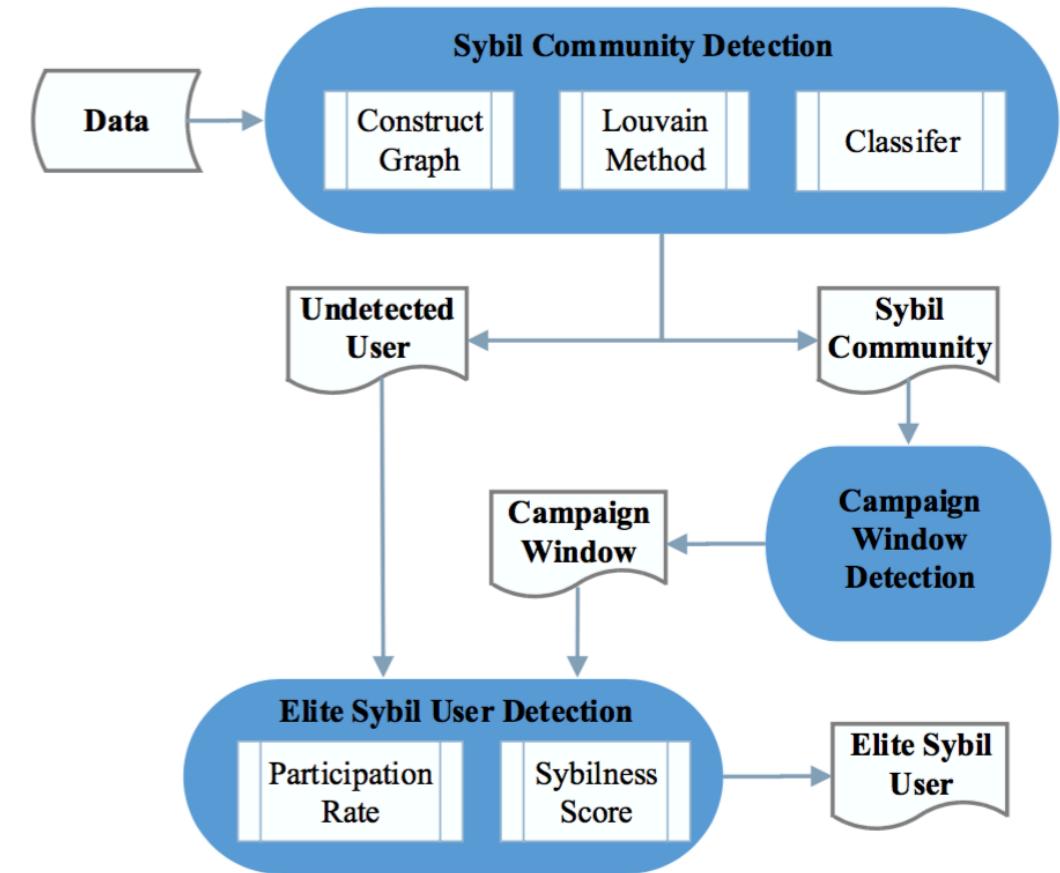
Sybil Community Detection



Campaign Window Detection



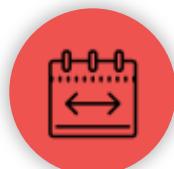
Elite Sybil User Detection



Overview



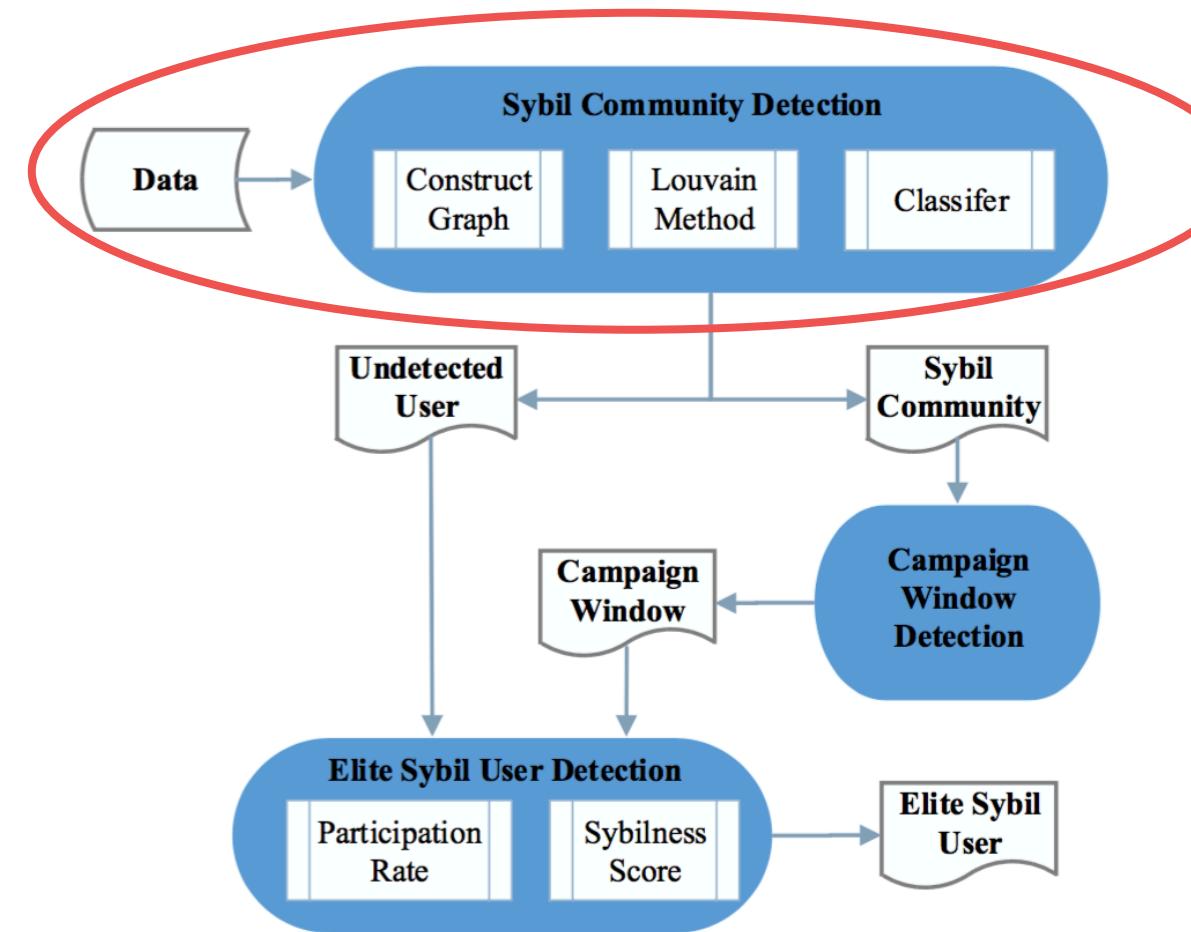
Sybil Community Detection



Campaign Window Detection



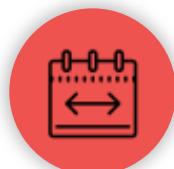
Elite Sybil User Detection





Sybil Community Detection

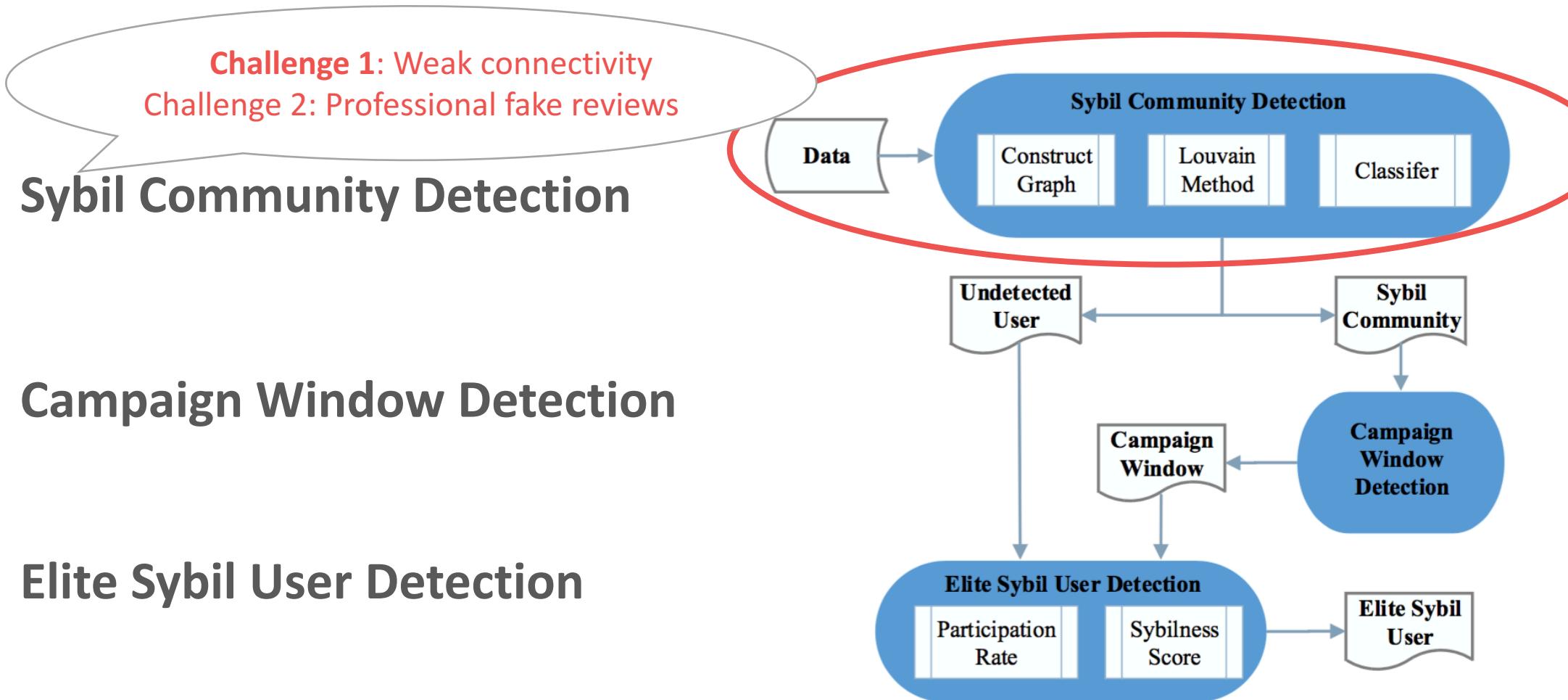
Challenge 1: Weak connectivity
Challenge 2: Professional fake reviews



Campaign Window Detection



Elite Sybil User Detection



Challenge 1: Build the relation between users via collusive reviews.



Sybil Community Detection

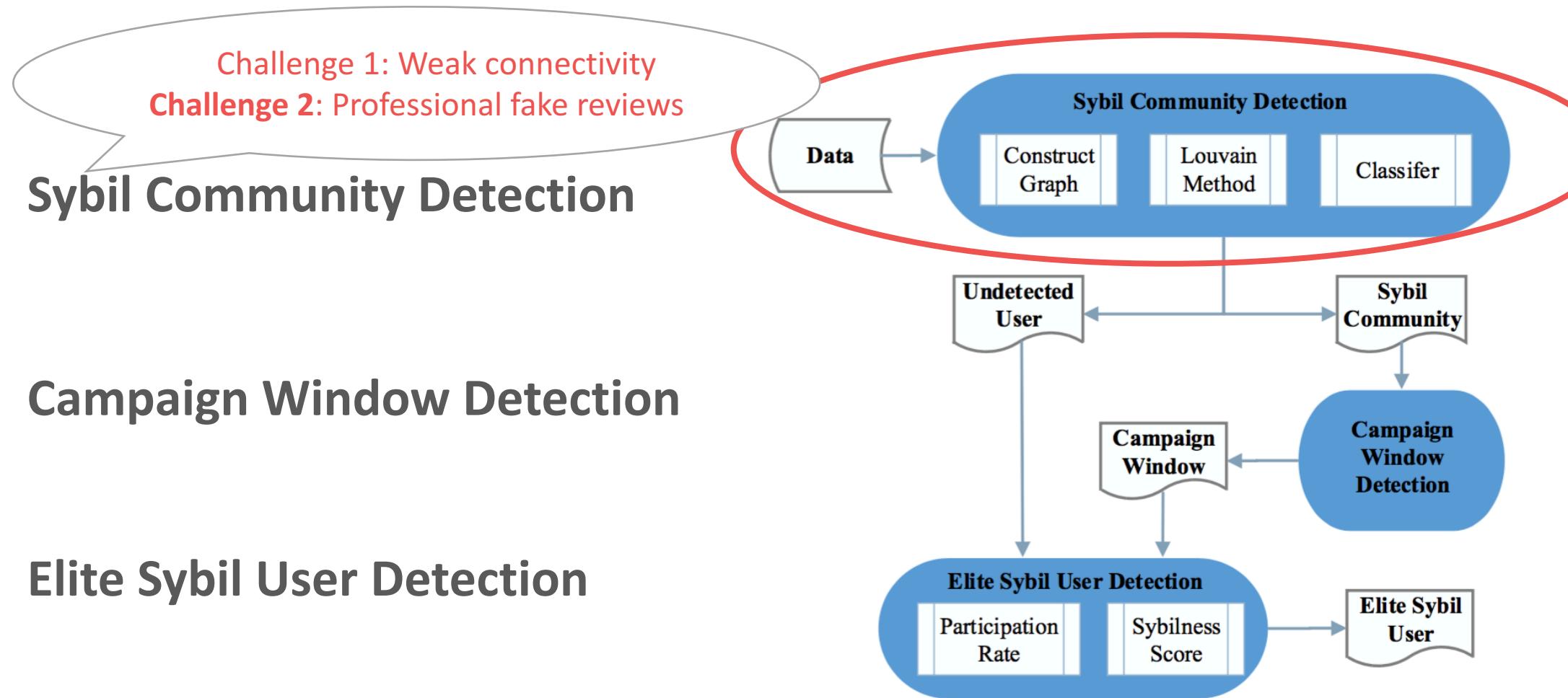
Challenge 1: Weak connectivity
Challenge 2: Professional fake reviews



Campaign Window Detection



Elite Sybil User Detection



Challenge 2: Use features irrelevant to user profiles and review content.
 (community-based features, network features, content-free user features)



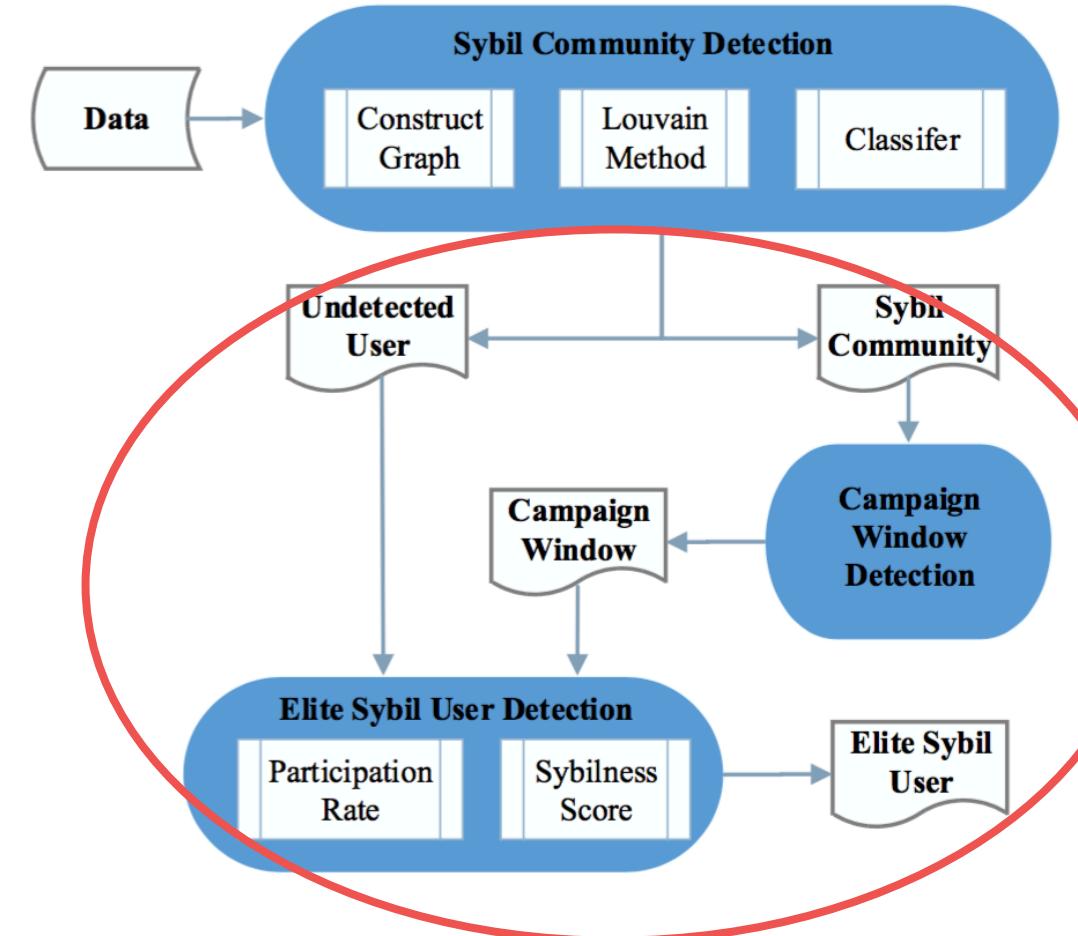
Sybil Community Detection



Campaign Window Detection

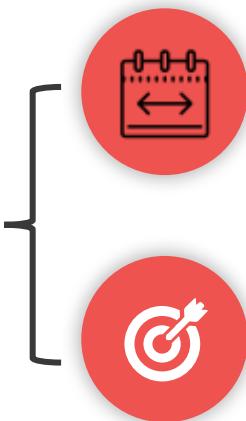


Elite Sybil User Detection



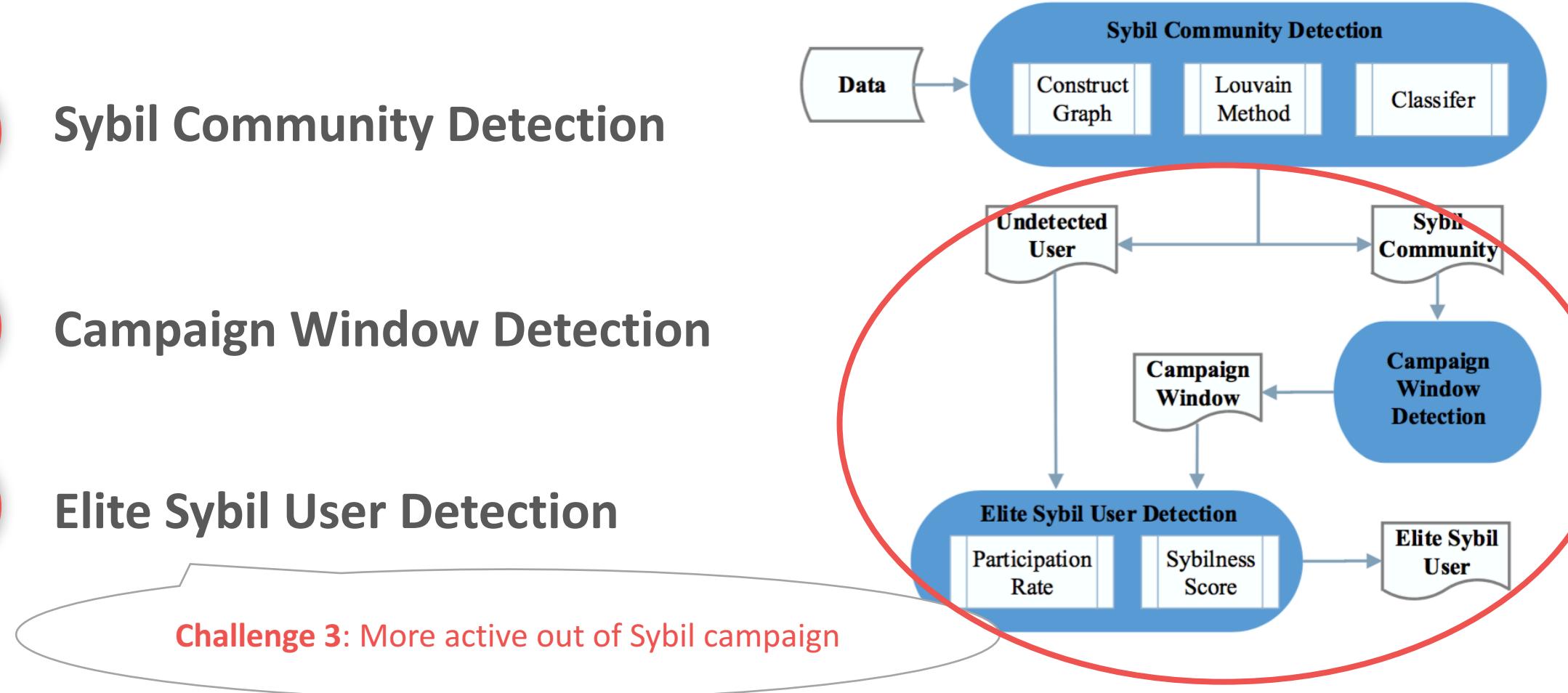


Sybil Community Detection



Campaign Window Detection

Elite Sybil User Detection



Challenge 3: Use campaign windows to filter out the reviews irrelevant to Sybil Campaigns

02

ELISEDET

Build Relations between Users

01

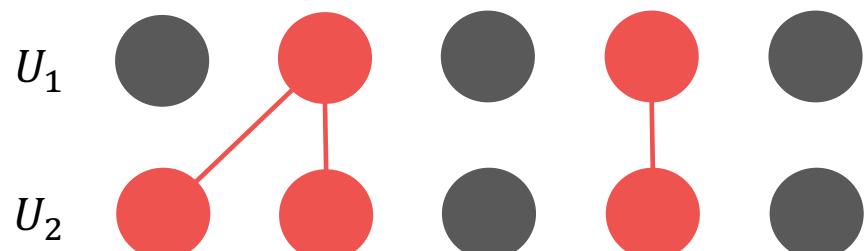
Collusive reviews

Posted by two users in a same store in a short time period ΔT with the same rating (same goal).

02

Collusive relation between users

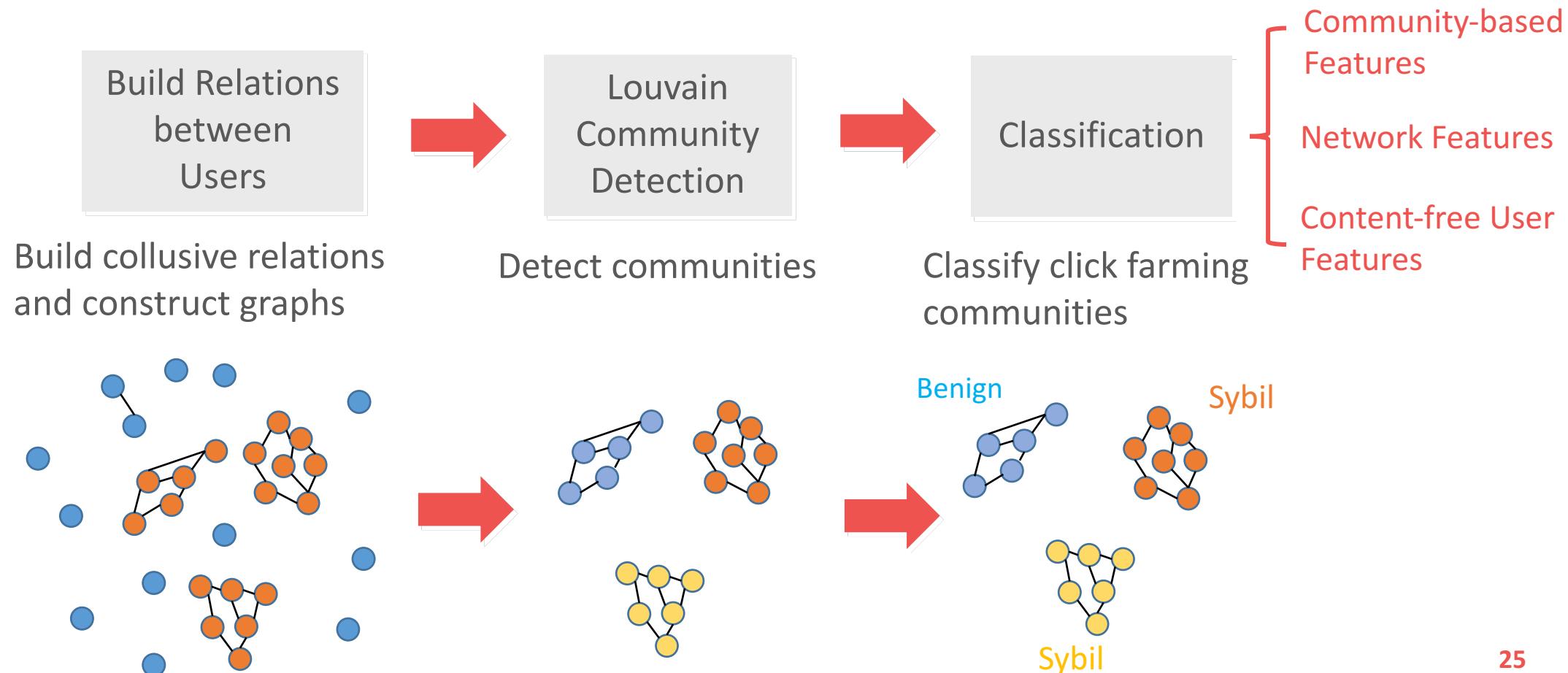
$$R(U_1, U_2) = \frac{|\text{Collusive Reviews}|}{|\text{Reviews of } U_1| + |\text{Reviews of } U_2|}$$



$$R(U_1, U_2) = \frac{5}{5 + 5} = 0.5$$

Sybil Community Detection

Detect Sybil communities in three steps:



01 Sparse interval:

The number of weeks with at least one review is less than the number of weeks without any reviews.

02 Iteratively find and delete sparse review intervals

ALGORITHM 1: Detecting Campaign Time Windows

Input: A list L_{review} whose item $L_{review}[i]$ denotes the number of reviews posted in the i th week.

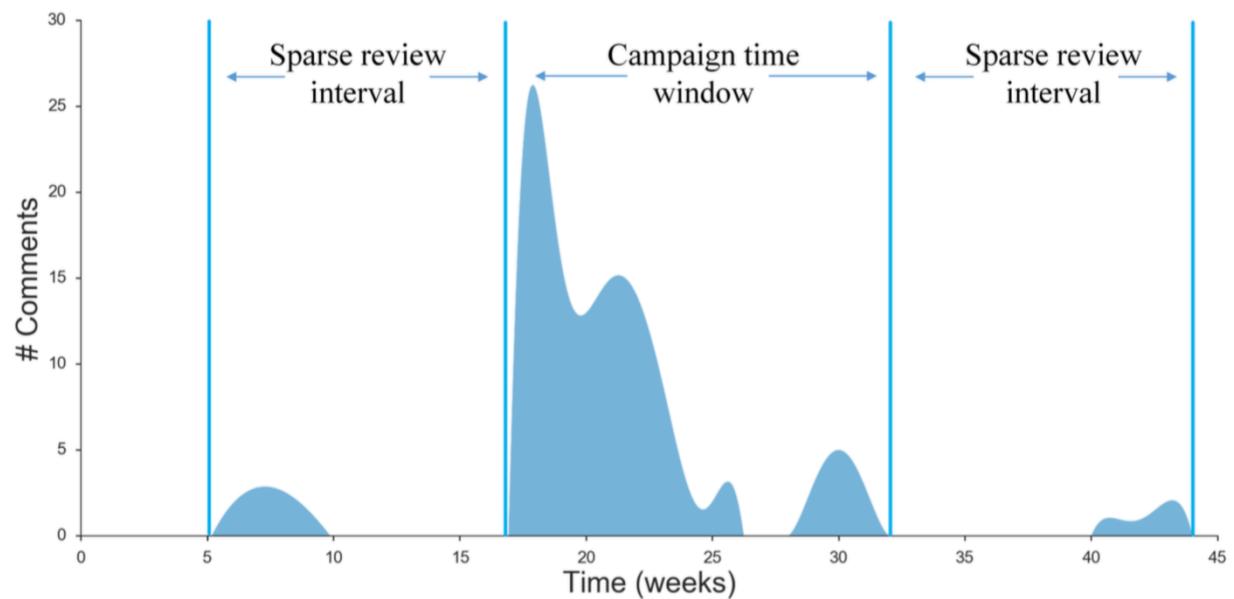
Output: The start point l and end point r of the campaign time window.

Initial:

```

1:  $l \leftarrow 0;$ 
2:  $r \leftarrow \text{length}(L_{review}) - 1;$ 
3: while (true) do
4:    $I_{l,l'} \leftarrow \text{find(left}, l);$  {Find the first sparse interval  $I_{l,l'}$  from left.}
5:    $I_{r',r} \leftarrow \text{find(right}, r);$  {Find the first sparse interval  $I_{r',r}$  from right.}
6:   if ( $l' = r$  and  $r' = l$ ) then {There is no sparse interval.}
7:     break;
8:   end if
9:   if ( $|I_{l,l'}| \leq |I_{r',r}|$ ) then {Choose the interval with fewer reviews.}
10:     $l \leftarrow l' + 1;$ 
11:  else
12:     $r \leftarrow r' - 1;$ 
13:  end if
14: end while
15: return  $l, r;$ 

```



02

ELISEDET

Elite Sybil User Detection

01

Participation rate between users and communities

Just consider the reviews posted in Sybil campaigns.

$N_C(k)$: the accumulated number of reviews posted C within the kth time window of community C.

N_C^{max} : the maximum number of reviews posted within all time windows of community C.

$$P_C(k) = \frac{N_C(k)}{N_C^{max}} \quad \text{Indicate the importance of kth time interval}$$

$$N_{u \in C} = \sum_k P_C(k) \cdot N_{u \in C}(k) \quad \text{Indicate the importance of reviews u posted for community C}$$

$$\text{Participation rate: } \rho_{u \in C} = \frac{1}{1 + \exp^{-\frac{N_{u \in C} - \mu_C}{\sigma_C}}}, \quad \text{for any } u \in C, \quad \text{Standardize with sigmoid function}$$

02

Sybilness

$$f(u) = \sum_C \rho_{u \in C} \cdot N_{u \in C}$$

Perceived likelihood indicating if a user is an elite Sybil user

Outlines

01

Elite Sybil Attacks

02

ELISEDET

03

Evaluation & Measurement

04

Conclusion

Detection Results

Dataset

32,933 stores, 3,555,154 users and **10,541,931 reviews**



Community

566 Sybil communities with 22,324 users and 144 benign communities with 5,222 users

Elite Sybil User

1. User u does not belong to any Sybil communities.
2. There is a community C and the user participation rate $\rho_{u \in C}$ is larger than 0.5 (average participation rate in a community).
3. **12,292 elite Sybil users.**

Detection Results

Dataset

32,933 stores, 3,555,154 users and **10,541,931 reviews**



Community

566 Sybil communities with 22,324 users and 144 benign communities with 5,222 users

Elite Sybil User

1. User u does not belong to any Sybil communities.
2. There is a community C and the user participation rate $\rho_{u \in C}$ is larger than 0.5 (average participation rate in a community).
3. **12,292 elite Sybil users.**

TABLE III. BREAKDOWNS OF STORES

Type	# Stores	# Overhyped Stores	Percentage of Overhyped Stores
Cinema	235	71	30.21%
Hotel	1,738	134	7.71 %
Restaurant	22,474	1,244	5.54 %
Entertainment	1,384	73	5.27 %
Wedding Service	320	8	2.50 %
Beauty Store	1,460	35	2.40 %

Sybil Community

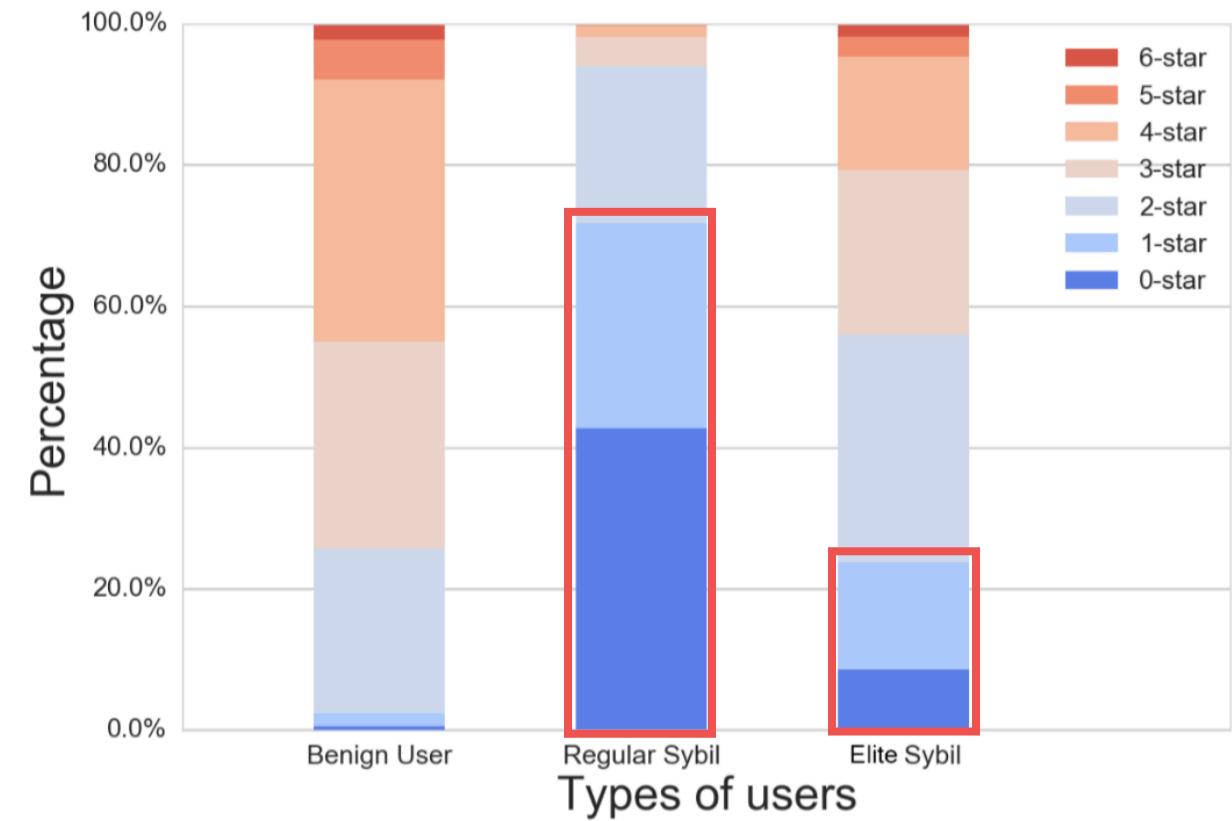
Classifier	Precision	Recall	F1	AUC
<i>Decision tree</i>	93.80 %	92.90 %	93.60 %	92.83 %
SVM	96.74 %	96.47 %	96.45%	99.42%
<i>GNB</i>	94.21 %	93.44 %	93.57 %	97.64 %
<i>KNN</i>	96.75 %	96.47 %	96.50 %	97.45 %
<i>Ada boost</i>	93.84 %	93.54 %	93.60 %	97.92 %
<i>Random forest</i>	93.16 %	94.01 %	92.99 %	97.42 %

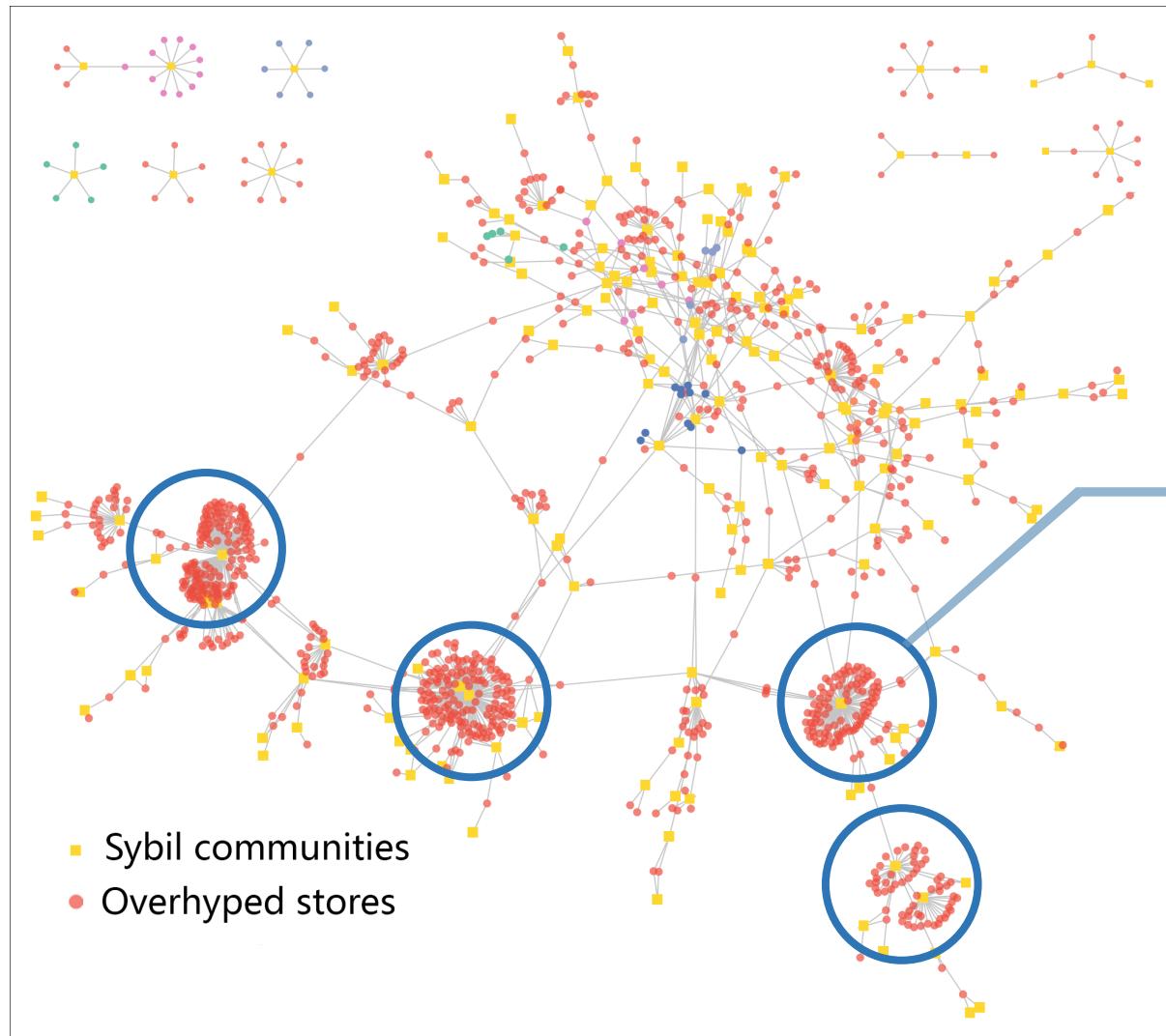
Elite Sybil User

Type of users	Precision
Sybilness top 1,000 users	93.80%
Random 1,000 users	90.70%

User-level star ratings

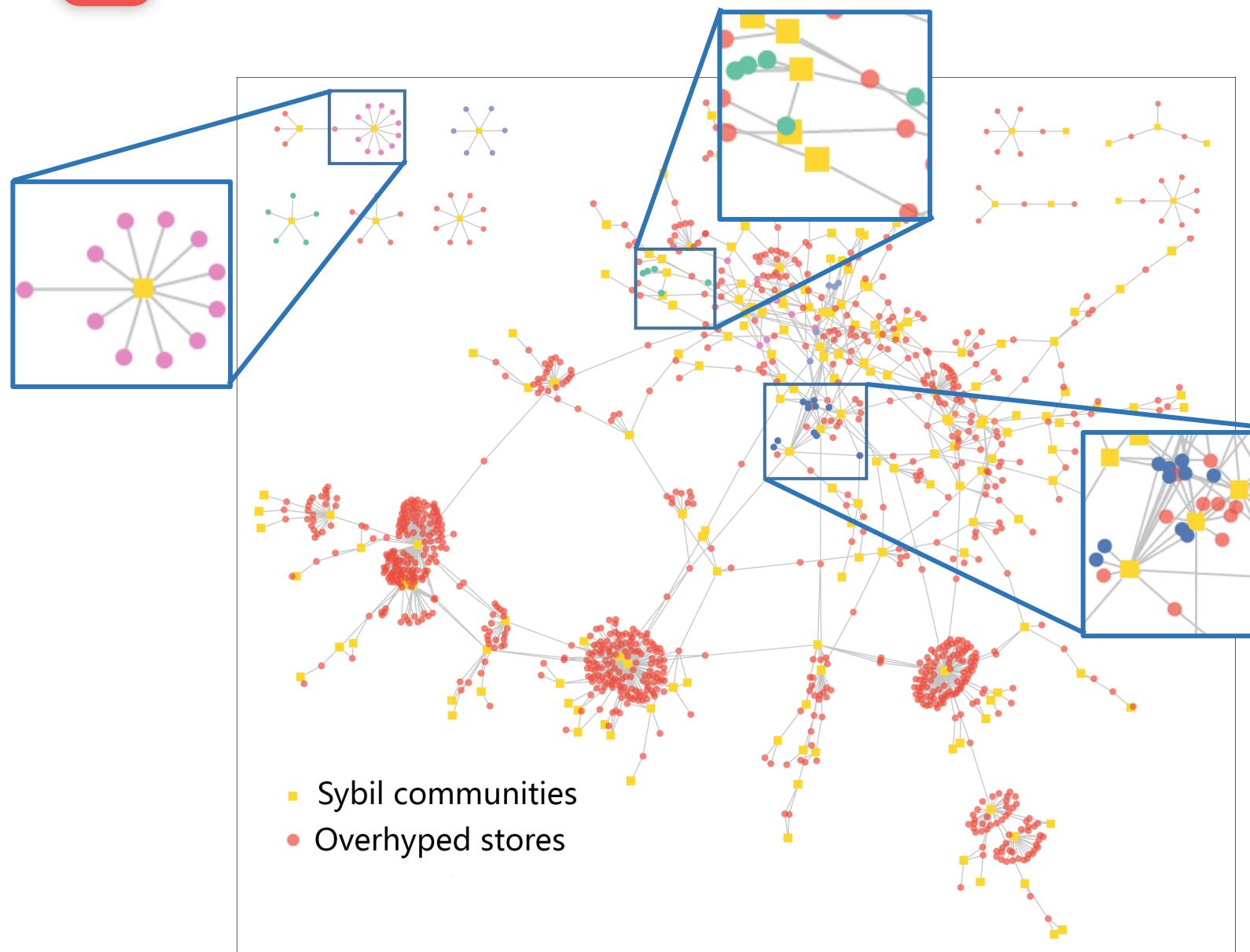
Compared to regular Sybil users,
elite Sybil users have higher user levels.





Sybil Communities

Some Sybil communities provide services for a huge number of overhyped stores.



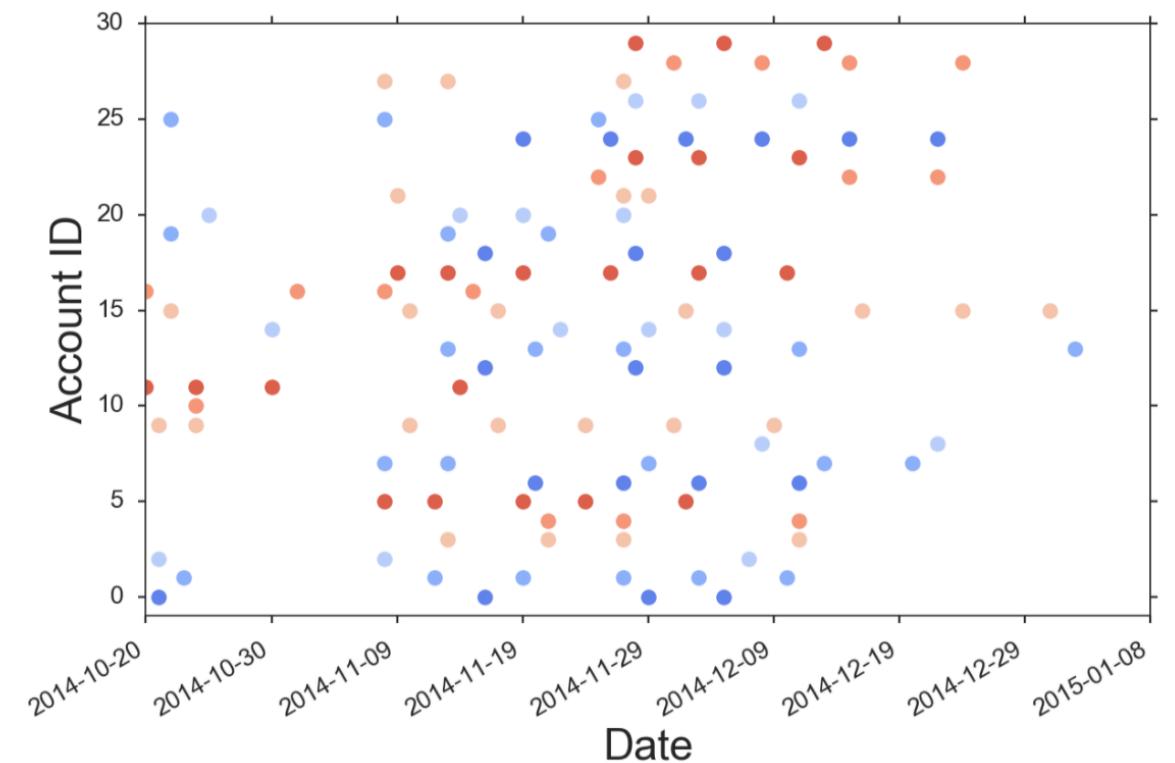
Chain store

1. Except for red dots, other dots with same colors are chain stores.
2. 12.37% of Sybil communities post fake reviews for chain stores.

User posting period

Elite Sybil users post fake reviews periodically.
Evading detection method basing on posting frequency.

Case Study



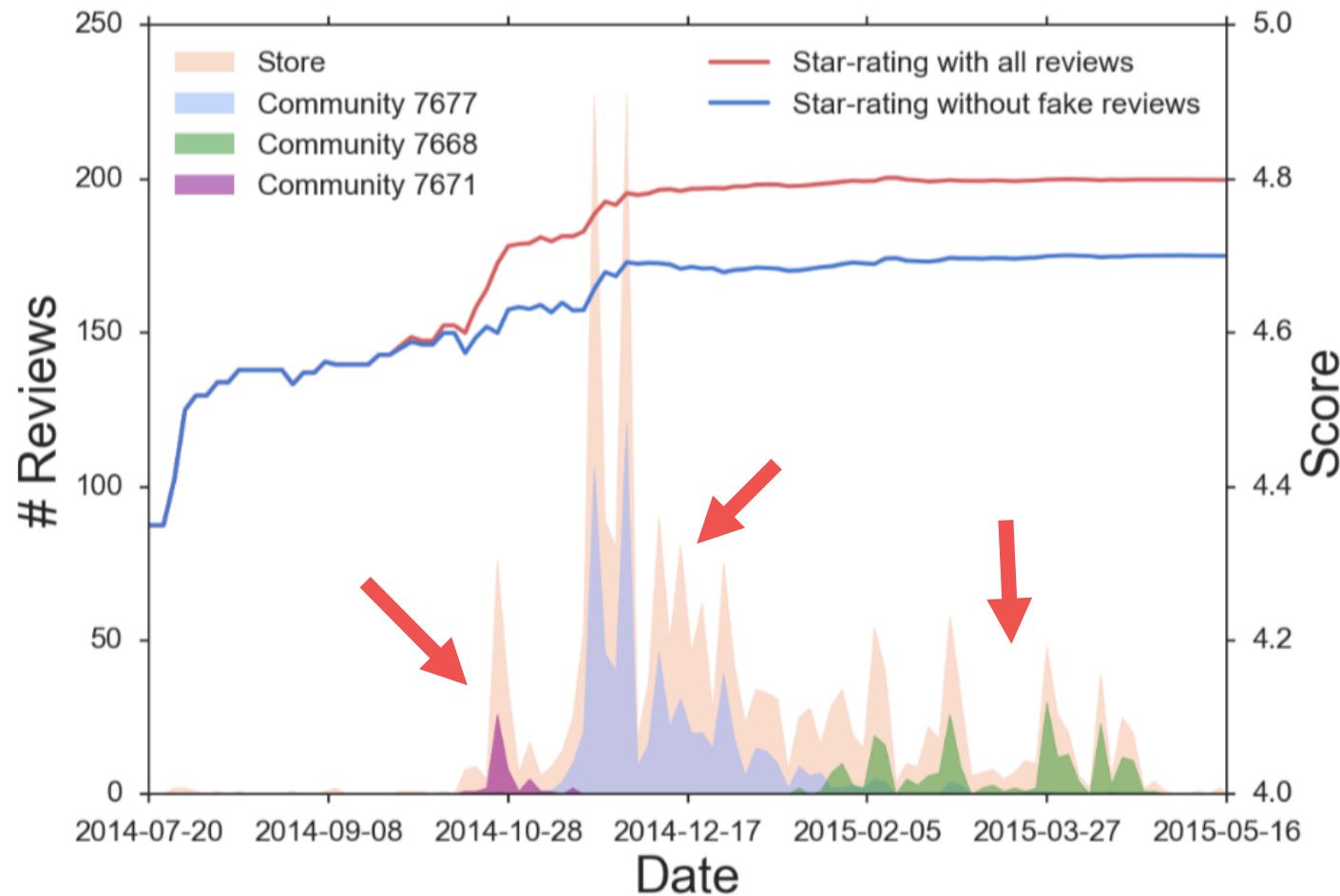
Reviews posted by Community 4559 in Store 4112200

Case Study

The review spike

Three different Sybil communities posted fake reviews in a store and caused the review spikes.

Increase in the number of reviews.

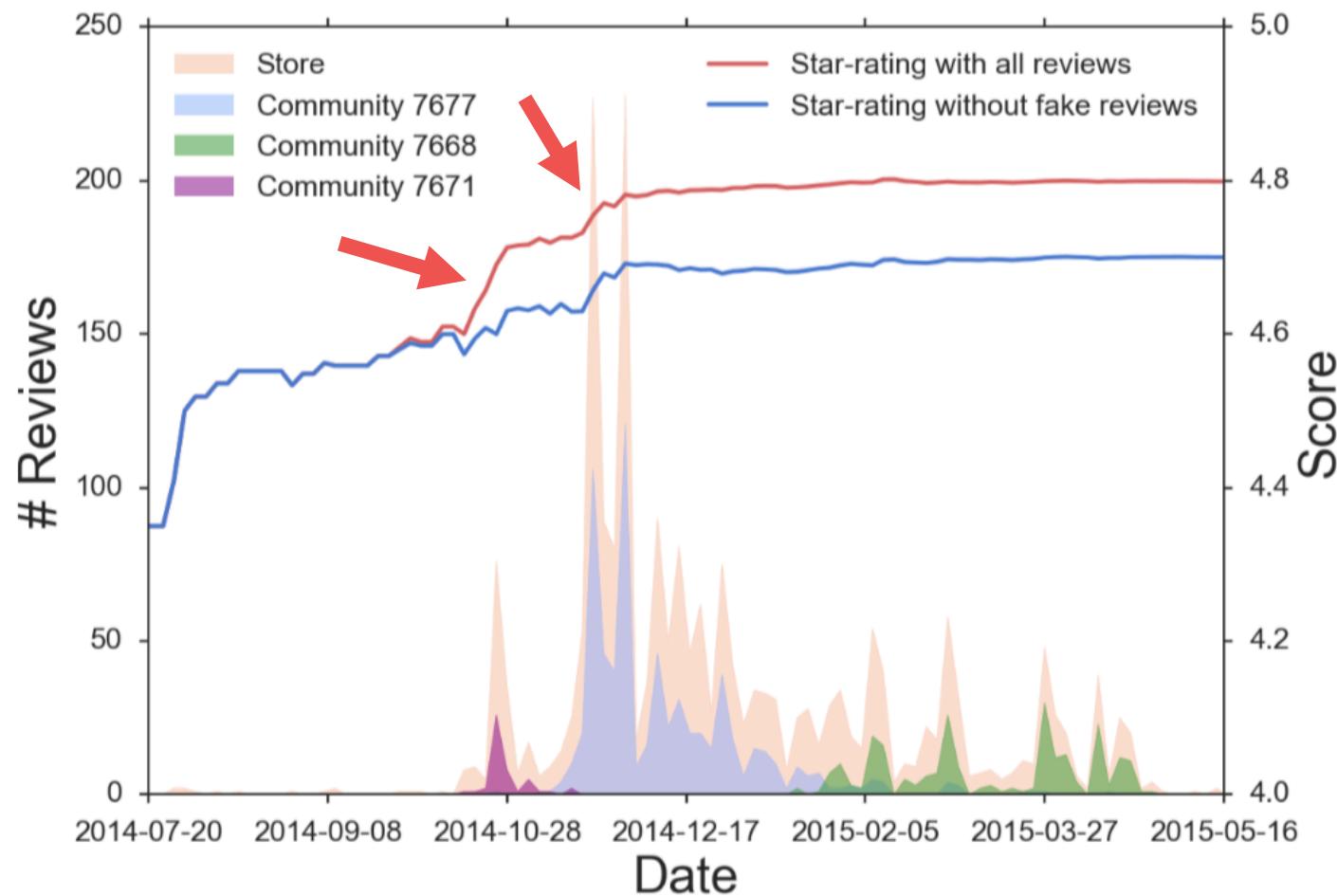


Variation of star ratings and the number of reviews of a hotel

Case Study

Increase in rating

The reviews spikes lead to increases in the rating of the store.



Variation of star ratings and the number of reviews of a hotel

Outlines

01

Elite Sybil Attacks

02

ELISEDET

03

Evaluation & Measurement

04

Conclusion

Novel attack: Elite Sybil attack

Elite Sybil users are more spread out temporally, craft better edited contents, but have fewer reviews filtered

Elite Sybil Detection System: ELISEDET

Both highly effective and scalable as a standalone system

Measurement about elite Sybil users

Elite Sybil have higher user-levels and deliberately control the posting time



Thanks for listening!

Questions?

sjtu.zhenghaizhong@gmail.com